SANDIA REPORT
SAND2018-9197
Unlimited Release
Printed August 2018

Evolution of Water Narratives in Local US Newspapers: A Case Study of Utah and Georgia

Thushara Gunda, Matthew D. Sweitzer, Kevin T. Comer, Conor Finn, Sandra Murillo-Sandoval, and Jennifer Huff

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology and Engineering Solutions of Sandia, LLC.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone:	(865) 576-8401
Facsimile:	(865) 576-5728
E-Mail:	reports@osti.gov
Online ordering:	http://www.osti.gov/scitech

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone:	(800) 553-6847
Facsimile:	(703) 605-6900
E-Mail:	orders@ntis.gov
Online order:	https://classic.ntis.gov/help/order-methods/

SAND2018-9197
Printed August 2018
Unlimited Release

Evolution of Water Narratives in Local US Newspapers: A Case Study of Utah and Georgia

Thushara Gunda[1,+], Matthew D. Sweitzer[2], Kevin T. Comer[3], Conor Finn[4,5], Sandra Murillo-Sandoval[6], and Jennifer Huff[7]

[1]Sandia National Laboratories, Albuquerque, NM 87185, United States
[2]School of Communication, Ohio State University, Columbus, OH 43210, United States
[3]MITRE Corporation, McLean, VA 22102, United States
[4]Centre for Complex Systems, Faculty of Engineering & IT, The University of Sydney, NSW 2006, Australia
[5]CSIRO Data 61, Marsfield, NSW 2122, Australia
[6]National Polytechnic Institute (IPN), Mexico City 07738, Mexico
[7]Burke Museum, Seattle, WA 98195, United States
[+] Corresponding author: tgunda@sandia.gov, Sandia National Laboratories, P. O. Box 5800, Albuquerque, New Mexico 87185-MS1137

Abstract

Narratives about water resources have evolved, transitioning from a sole focus on physical and biological dimensions to incorporate social dynamics. Recently, the importance of understanding the visibility of water resources through media coverage has gained attention. This study leverages recent advancements in natural language processing (NLP) methods to characterize and understand patterns in water narratives, specifically in 4 local newspapers in Utah and Georgia. Analysis of the corpus identified coherent topics on a variety of water resources issues, including weather and pollution. Closer inspection of the topics revealed temporal and spatial variations in coverage, with a topic on hurricanes exhibiting cyclical patterns whereas a topic on tribal issues showed coverage predominantly in the western newspapers. We also analyzed the dataset for sentiments, identifying similar categories of words on trust and fear emerging in the narratives across newspaper sources. An analysis of novelty, transience, and resonance using Kullback-Leibler Divergence techniques revealed that topics with high novelty generally contained high transience and marginally high resonance over time. Although additional analysis needs to be conducted, the methods explored in this analysis demonstrate the potential of NLP methods to characterize water narratives in media coverage.

ACKNOWLEDGMENTS

This report summarizes the efforts undertaken by the team as part of the Santa Fe Institute (SFI)'s Complex Systems Summer School in 2018. So the authors would like to acknowledge the various lecturers, attendees, and affiliates of SFI that helped inform this effort, including Marion Dumas and Simon DeDeo. The authors would also like to thank Jonathan Gilligan, Allison Witte, and George Hornberger at Vanderbilt University, who helped conceptualize and download the initial corpus used in this analysis.

TABLE OF CONTENTS

FIGURES

TABLES

NOMENCLATURE

Abbreviation	Definition
BLM	Bureau of Land Management
FREX	FRequently and EXclusively
GA	Georgia
GRU	gated recurrent unit
KLD	Kullback-Leibler Divergence
LIWC	Linguistic Inquiry and Word Count
NLP	natural language processing
STM	structural topical modeling
UT	Utah

1. INTRODUCTION

Narratives about water resources have evolved, from focusing solely on physical and biological dimensions to also consider social dynamics (Gunda et al., 2017). System-oriented perspectives of water resource management as a social ecological system, in particular, highlight the complex nature of this resource (Edalat et al., 2018). Various dynamics influence the feedbacks between the social and physical aspects of the system, including food and energy production, recreation preferences, local rituals, and policies.

Recently, the “visibility” of water resources and subsequent influence on water-related behaviors has been highlighted by researchers (Brown, 2017; Quesnel & Ajami, 2017). Water resources can be made visible directly to consumers through physical proximity to reservoirs and wetlands as well as indirectly through local discourses, by public officials and media narratives (Brown, 2017). Evaluating media narratives, in particular, can provide a lot of insight into societal values because news coverage can not only influence but also reflect public opinion (Rogers & Dearing, 1988; Wei et al., 2017). Incorporating the influence of media narratives into system dynamics models, for example, has helped to explain rebound behaviors in water consumption post-droughts in California (Gonzales & Ajami, 2017). Quality of media coverage is also important to consider since media narratives help construct a common understanding of water issues (Verón, 1997).

Analysis of newspaper coverage on water resources has been undertaken by numerous researchers, including Quesnel & Ajami (2017) in California, Treuer et al., (2017) in Miami-Dade, Wei et al., (2017) in Australia, and Wu et al., (2018) in China. These studies highlight the influence of media coverage on actual water consumption, trends in economic vs environmental framing, and generally negative tone of news coverage (Quesnel & Ajami., 2017; Treuer, et al., 2017; Wei, et al., 2017; Wu, et al., 2018). This study builds on these previous efforts by leveraging recent advancements in natural language processing (NLP) methods to characterize and understand patterns in coverage of water issues over time and space. Our expectation is that the techniques employed have high potential for providing invaluable insights into local variations in the public discourse of water.

2. METHODS

For this analysis, we take advantage of a downloaded corpus of newspaper articles by Jonathan Gilligan and the Water Conservation group (including Allison Witte and George Hornberger) at Vanderbilt University. The corpus encompasses any article containing the word ‘water’ from 1990 to late 2017 in any of 37 local newspapers across 34 states in the United States contained in the LexisNexis database (Table 1). The full corpus contains ~1.8M articles but the number of articles for each newspaper varies, influenced in part by the different coverage periods within LexisNexis (Table 1; Figure 1). This initial dataset was converted from the original xml files into data frames with relevant information (e.g., headline, byline, content, date, source, and section). Additional processing conducted on the dataset included cleaning up sources and removing duplicate entries.

There are various aspects of water covered in the media. Structural topic modeling (STM) can help us group articles into those covering similar topics to help us refine our full dataset into those of most interest (Roberts, et al., 2014; Roberts et al., 2016; see also http://www.structuraltopicmodel.com). Specifically, STM groups documents using FREX words (i.e., words that are occurring FRequently and EXclusively within a group). STM allows use of covariates to account for the way topic coverage may vary systematically as a function of some document characteristic. For example, we may expect that the topics covered by one newspaper will be generally different from the topics covered by another. By including the articles’ sources as a covariate in the model estimation, we can account for this underlying structure and the connection a source serves between different articles.

We employ STM in two ways, one as a way to filter the full corpus to articles actually pertaining to water resources and another as a way to group the subset corpus into coherent topics for further analysis. Using topic models to filter documents of interest (Gao et al., 2013) is particularly useful for our corpus because metaphorical uses of “water” were present in the initial corpus. Once the corpus was filtered to the relevant articles, STM was executed again to identify coherent topics of interest. The number of topics used by STMs – like other forms of topic modeling – is user-specified and thus, can be subjective. To address this, the stm package in R (Roberts et al., 2018), includes a “searchK” function, which calculates the degree to which the corpus can be cleanly separated at various levels of k (number of topics). Due to time constraints, the initial execution of the STM for filtering the documents was conducted with 100 topics while the second execution of the STM for filtering the topics was informed by searchK results. For the initial filtering of documents, any article that contained the identified topics of interest in its top 5 topic fits was included in the second set of topics analysis. Both executions of STM considered covariates of the article source (i.e., newspaper name) and date of the article.

The resulting topics were analyzed with a variety of mathematical approaches to understand patterns in the narrative. For example, we looked at variations in topic coverage as a function of time and space, using the continuous and point estimate methods in the “estimateEffects” function in stm respectively (Roberts et al., 2018). We performed sentiment analysis to understand the general tone of the topics using the tidytext package in R (Silge & Robinson, 2016); net sentiment scores were estimated using the Bing lexicon, which classifies words into binary (positive or negative) scores while sentiment variations by source were evaluated using the NRC lexicon, which classifies words into emotion categories such as fear, anticipation, joy, and surprise among others. We also looked at relationships between topics using hierarchical clustering analysis, networks of correlations between topics, and Spinglass community detection methods (de Arruda et al., 2016; Farrell, 2015).

Lastly, following Barron et al., (2018), we used Kullback-Leibler Divergence (KLD) to assess the "surprise" or deviations in topic patterns. Specifically, we analyzed the dataset for novelty (defined as the "new-ness" or surprise of ideas given what you know about the past ideas), transience (defined as surprise of ideas in the present given what you know about the future), and resonance (defined as the difference between novelty and transience to capture how topics from the past stick around into the future) by adapting the python code supplied by Barron et al., (2018) (see also https://github.com/CogentMentat/NTRexample_FRevNCA). The KLD analysis was conducted for the full corpus (i.e., average of all topics) as well as for specific topics of interest using a pointwise approach. A window length *w* dictates the number of articles previous and after the current article evaluated for novelty and resonance respectively. All analyses were conducted in R, except the KLD calculations, which were conducted in python.

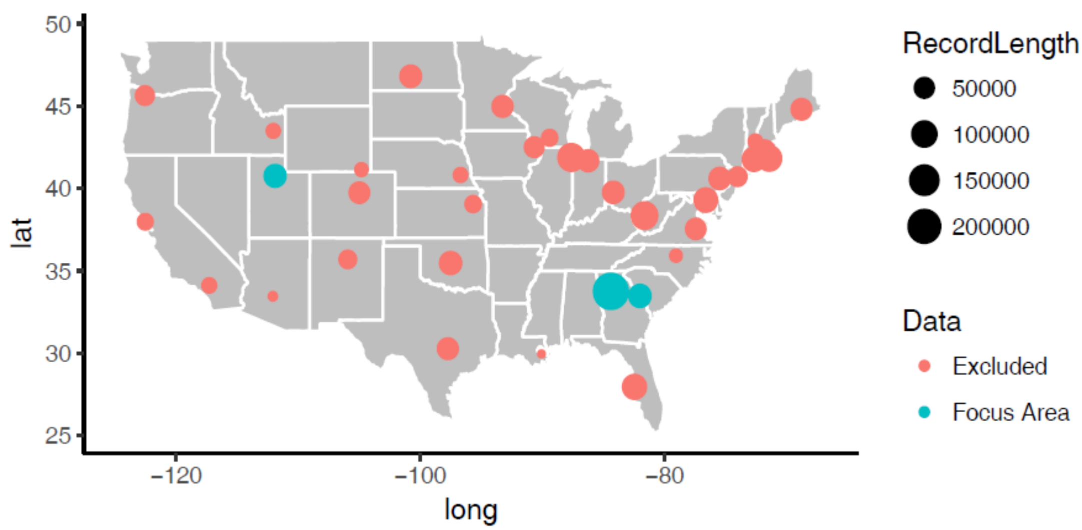

Figure 1. Areas of newspaper coverage in full corpus. Due to time constraints, this analysis focused on four newspapers within the corpus (blue dots), two each in Utah (Salt Lake Tribune and Deseret Morning News) and Georgia (Atlanta Journal Constitution and Augusta Chronicle).

Table 1. Coverage period and article counts of newspapers in full corpus.

Newspaper	Headquarters (City)	Coverage Begin	Coverage End	Article Count
Arizona Capitol Times	Phoenix, AZ	2-Jun-03	N/A	951
Bangor Daily News	Bangor, ME	28-Jul-94	N/A	54211
Brattleboro Reformer	Brattleboro, VT	13-Oct-04	N/A	14488
Chapel Hill Herald	Chapel Hill, NC	1-Jan-95	N/A	7958
Chicago Daily Herald	Chicago, IL	1-Jul-97	N/A	114782
Daily Journal of Commerce	Portland, OR	22-Feb-01	N/A	3289
Daily News	New York City, NY	1-Mar-95	N/A	39307
Dayton Daily News	Dayton, OH	1-Jan-94	N/A	60196
Deseret Morning News	**Salt Lake City, UT**	**1-Jan-96**	**N/A**	**66405**
Idaho Falls Post Register	Idaho Falls, ID	1-Jan-93	N/A	13760
Lincoln Journal Star	Lincoln, NE	13-Jun-96	N/A	14575
Marin Independent Journal	San Rafael, CA	20-Aug-02	N/A	21309
Providence Journal	Providence, RI	1-Jan-94	N/A	97616
Richmond Times Dispatch	Richmond, VA	1-Nov-95	N/A	51544
San Bernardino Sun	San Bernardino, CA	14-Sep-01	N/A	16040
South Bend Tribune	South Bend, IN	1-Jan-94	N/A	57332
Star Tribune	Minneapolis, MN	1-Sep-91	N/A	55871
Telegram & Gazette	Worcester, MA	14-Feb-89	19-May-15	61441
Telegraph Herald	Dubuque, IA	28-Aug-95	N/A	46415
The Atlanta Journal-Constitution	**Atlanta, GA**	**1-Jan-91**	**N/A**	**126674**
The Augusta Chronicle	**Augusta, GA**	**1-Jan-92**	**N/A**	**68606**
The Austin American-Statesman	Austin, TX	1-Jan-94	N/A	54397
The Baltimore Sun	Baltimore, MD	1-Jan-90	N/A	78410
The Bismarck Tribune	Bismarck, ND	1-Jan-93	N/A	60959
The Capital Times	Madison, WV	1-Jan-92	N/A	22917
The Charleston Gazette-Mail	Charleston, WV	1-Jan-94	N/A	116407
The Columbian	Vancouver, WA	26-May-94	N/A	39935
The Daily Oklahoman	Oklahoma City, OK	1-Jan-92	N/A	66649
The Denver Post	Denver, CO	1-Dec-93	N/A	52675
The Hartford Courant	Hartford, CT	1-Jun-91	N/A	76591
The Journal of Jefferson Parish	New Orleans, LA	4-Nov-05	24-May-08	161
The Morning Call	Allentown, PA	11-Apr-91	N/A	59479
The Salt Lake Tribune	**Salt Lake City, UT**	**1-Jan-94**	**N/A**	**41097**
The Santa Fe New Mexican	Santa Fe, NM	1-Jan-94	24-Apr-13	31617
The Tampa Tribune	Tampa, FL	2-Feb-90	N/A	91092
The Wyoming Tribune-Eagle	Cheyenne, WY	26-Jun-97	N/A	10454
Topeka Capital-Journal	Topeka, KS	1-Jan-98	N/A	24212

Note. The four newspapers used in this analysis are bolded.

3. RESULTS

Due to time constraints associated with the Complex Systems Summer School, we restricted our analysis to 4 newspapers, two from Georgia (the Atlanta Journal Constitution and the Augusta Chronicle) and two from Utah (the Salt Lake Tribune and the Deseret Morning News; Figure 1). The coverage period of these 4 newspapers is pretty similar, extending back to the 1990s (Figure 2). An STM of k=100 topics on this subset corpus of 302,782 articles helped us to isolate topics of interest, including topic 7 (which contained FREX words such as hurricanes and floods), topic 13 (sewage and treatment), topic 34 (hot, ice, and weather), and topic 38 (dam, lake, and corps). In addition to FREX words, network graphs (in which topics served as nodes and their correlations served as ties, with a cutoff of r=0.10) based on topic correlations were used to identify 17 topics that seemed most relevant to water resources (Figure 3). These 17 topics were used to filter the corpus of 302,782 articles into a corpus of 65,611 articles.

The searchK function on the subset corpus helped us identify 13 as the number of topics for subsequent STM analysis (Figure 4). A separate STM was then estimated on the subset dataset with 13 topics (Table 2). There was some correlation between the topics related to emergencies (such as fire and weather) and topics related to legislative issues (such as tribes, contamination, energy, and dams; Figure 5). Similar groupings were observed using Spinglass community detection and clustering methods (results not shown). For comparison, STMs were run without metadata, with metadata, and with prevalence options; all 3 executions resulted in similar topics, so the STM with prevalence was retained to enable exploration of metadata effects. For example, Topic 3 on hurricanes exhibits a cyclical pattern every few years, indicative of large hurricanes (such as Andrew, Katrina, Irene, Rita) that make national news (Figure 6). Coverage of Topic 7 on tribal issues, on the other hand, has increased coverage after 2011, reflecting disputes between the Ute Tribes and the Bureau of Land Management (BLM) regarding water rights in Utah (Figure 6). There are also spatial differences in topic coverage, with the Augusta Chronicle and Deseret Morning News covering hurricane issues more than their counterparts in the same states (Figure 7). Coverage of tribal issues is, unsurprisingly, more prevalent in Utah than in Georgia (Figure 8).

We also performed sentiment analysis, both over time as well as over the entire corpus. When plotted over time, most of the topics discussed earlier came up as decidedly negative, such as Topic 3 (hurricanes; Figure 9). However, Topic 7 (tribal issues) revealed a more balanced sentiment analysis, particularly in more recent years, representing more positive or neutral news coverage regarding tribal issues in Utah since the resolution of disputes between the Utes and BLM (Figure 9). A more in-depth analysis with the NRC lexicon identified trust and fear as the dominant categories of sentiments across the four sources (Figure 10). Using the Bing lexicon, we plotted the specific words associated with positive and negative sentiments across the subset corpus, with words like emergencies, drought, and damage dominating the negative sentiments (Figure 11). Although a helpful first step, there are limitations to the sentiment analysis based on this bag-of-words approach. For example, the word “lead” (which is listed as positive) could mean “to lead” or “lead in the water”, and "clean" could refer to "lack of clean water" (a negative reflection of the current situation; Figure 11).

Finally, KLD plots of the full corpus show that transience does not increase as rapidly as novelty (i.e., sub-linear with respect to y=x) while resonance generally increases with novelty at small time frequencies (Figure 12). As we increase the size of the window of the KLD calculations, the fitted lines approach y=x and y=0 for novelty-transience and novelty-resonance respectively, indicating that articles with high novelty also have high transience but generally low resonance over time (Figure 12). Using pointwise KLD, patterns in specific topics were also analyzed. For example, a plot of resonance over time for Topic 3 (hurricanes) indicates peaks coinciding with notable hurricanes, such as Hurricane Katrina in Aug 2005 (Figure 13).

Table 2. FREX words by topic

Topic	FREX words
1: Home Efficiency	thermostat, odd-even, even-numbered, watering, dulley, tankless, royalgreen, flapper, low-flow, heat-related
2: Recreation (water)	pwc, pwcs, watercraft, bui, pontoon, houseboat, outboard, regatta, rowing, stand-alone
3: Weather	forecasters, nino, landfall, typhoon, haiti, meteorologists, nagin, galveston, katrinas, kph
4: Dams	acre-feet, julander, cfs, chub, mussel, acrefoot, mussels, quagga, salinity, snwa
5: Fire	firefighters, blaze, arson, extinguished, smoldering, freitag, mckone, two-alarm, christchurch, ufa, lightning-caused
6: Household Maintenance	heloise, soap, stain, vinegar, stains, wax, -ounce, detergent, glue, fragrance
7: Tribal and Legislative Issues	suwa, trump, jurrius, leavitt, lawmakers, sitla, rep, tribe, tribal, legislation
8: Contamination	epd, arsenic, superfund, reheis, phosphorus, epds, coliform, fecal, deq, dhec
9: Energy	nuclear, reactors, reactor, ethanol, nrc, deepwater, plutonium, bps, vogtle, fracking
10: Property	deed, thence, borrower, grantor, plat, lender, thereof, indebtedness, hereafter, augusta-richmond
11: Recreation (land)	hiking, hikers, trailhead, solitude, bikers, horseback, snowbasin, tallulah, snowmaking, lodges
12: City Code	ordinance, rezoning, annexation, rezone, fluoridation, byrne, splost, rezoned, rezonings, lathem
13: Education	leed, students, hookah, sloc, shes, teachers, student, arrington, teacher, curriculum

Figure 2. Number of articles in the UT and GA subset over time.

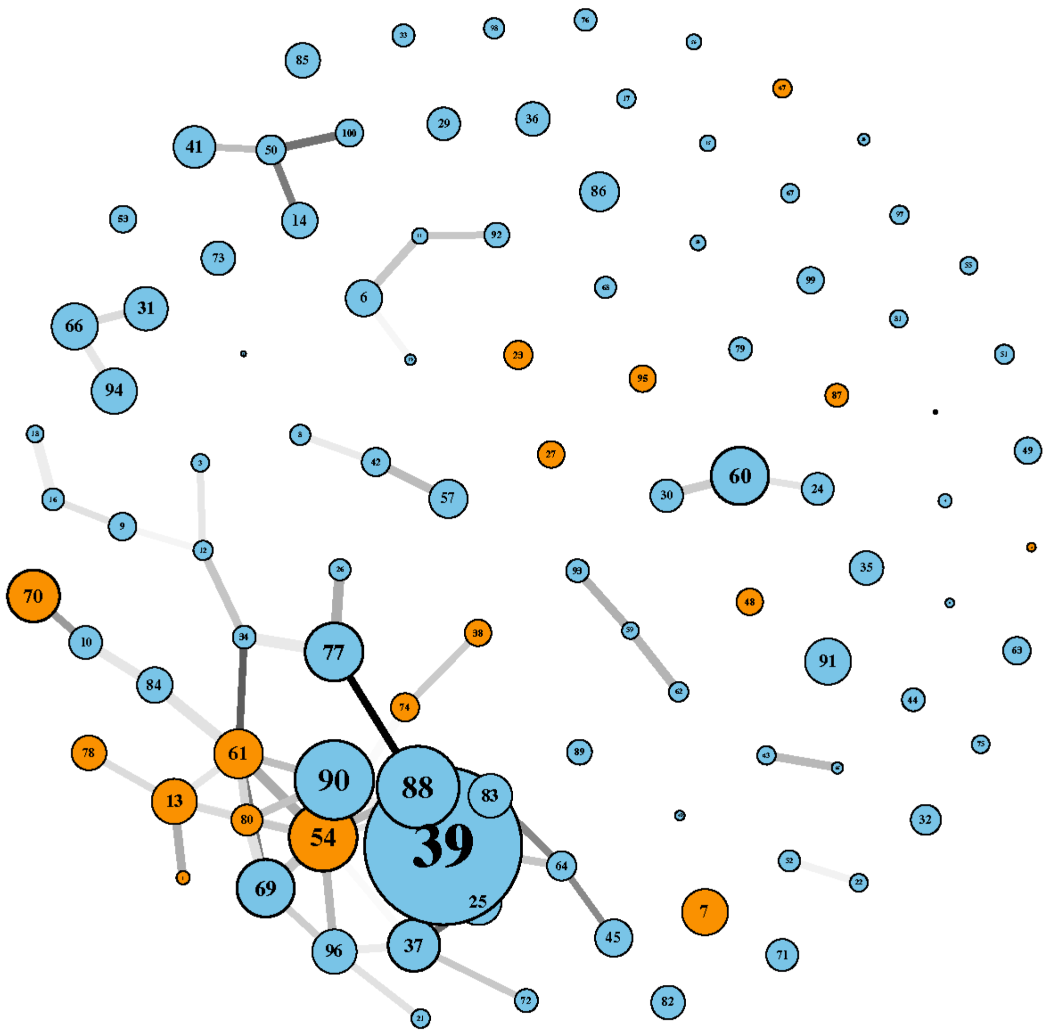

Figure 3. Network graph of initial 100 topic correlations. Node size corresponds to the proportion of the total corpus which includes that topic. Edges indicate positive correlations, with stronger correlations indicated by deeper color. Orange nodes are those topics which were included in the subsequent analysis.

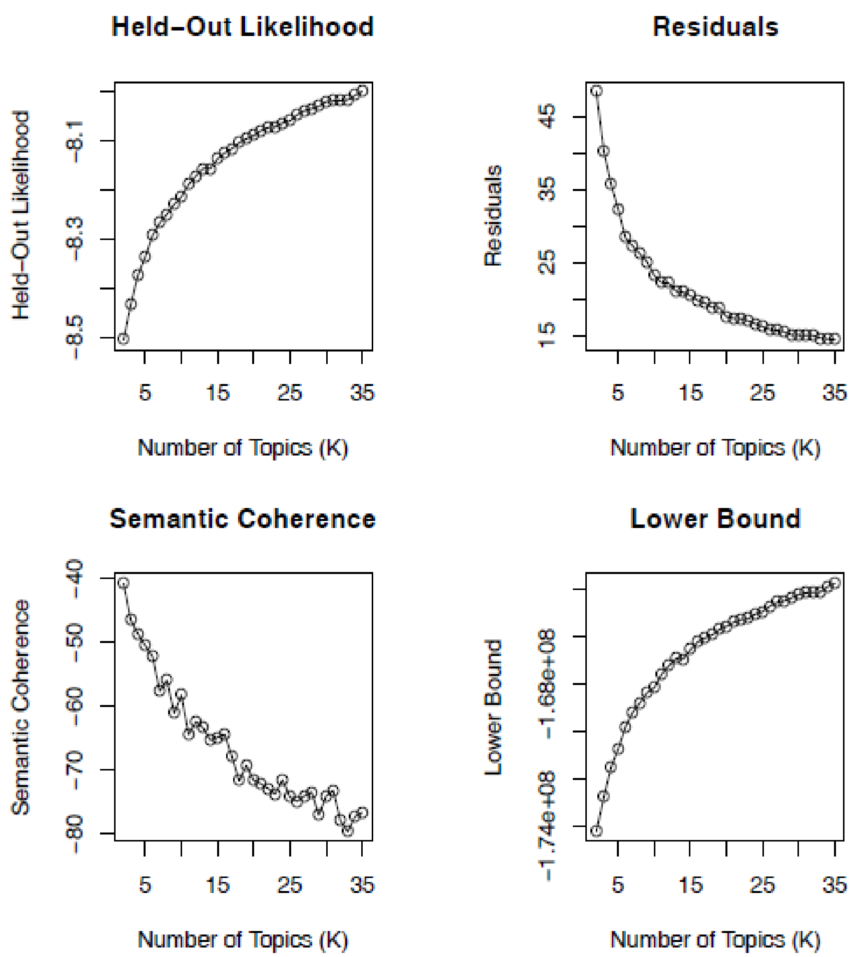

Figure 4. searchK results on the UT and GA subset. There is a small increase in semantic coherence at K=13 topics.

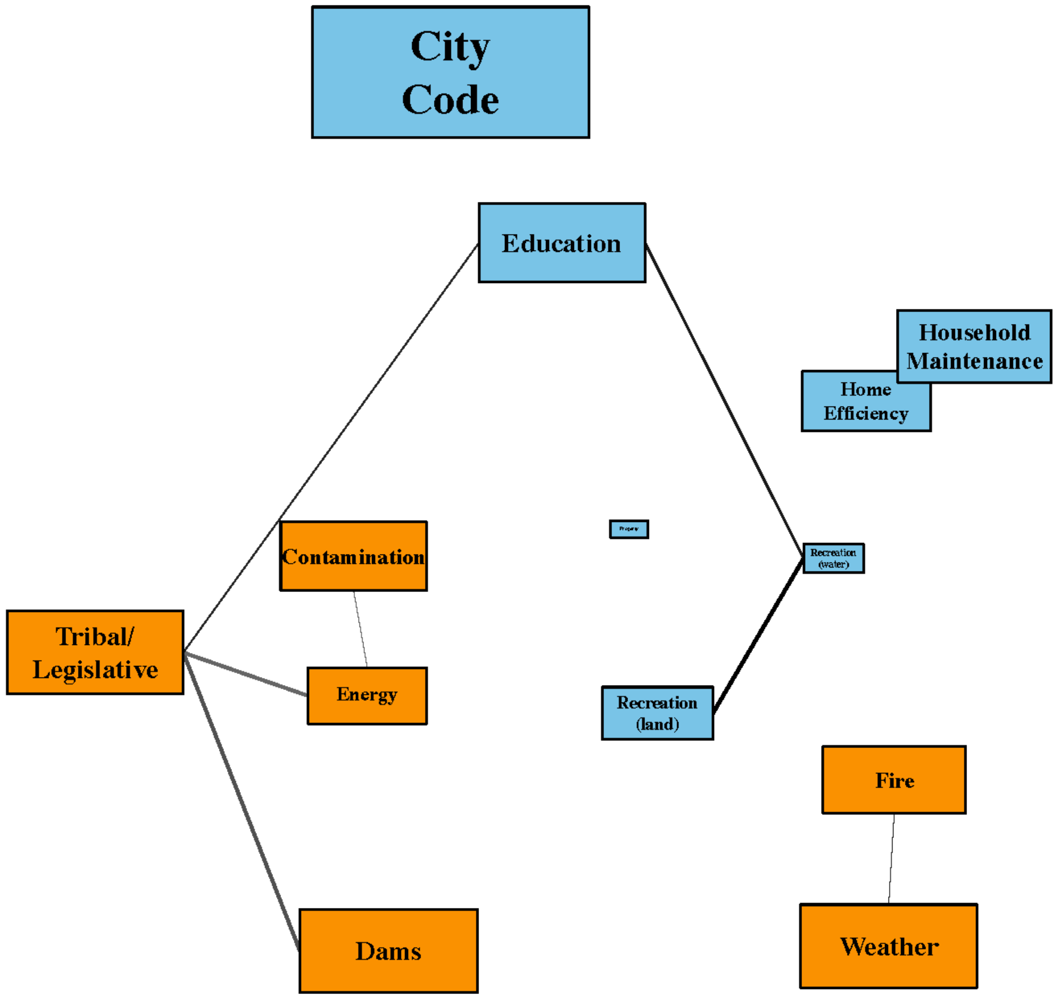

Figure 5. Network graph of final 13 topic correlations. Topics of interest within the final structural topic model are colored in orange. Line thickness indicates the strength of correlations.

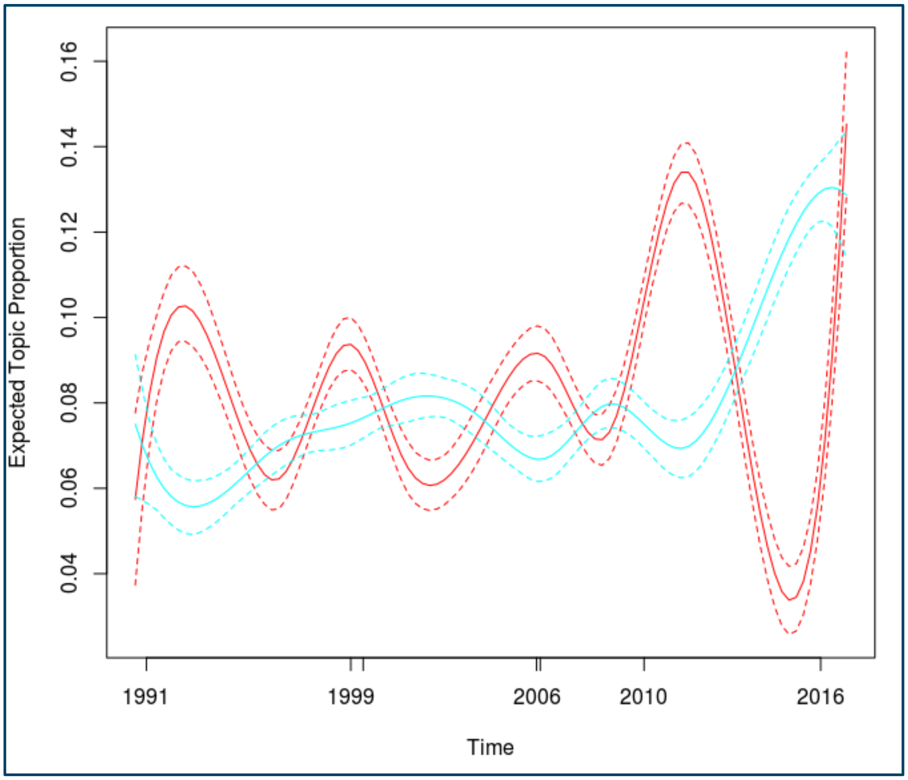

Figure 6. Coverage in topics 3 and 7 over time. Coverage of hurricane issues (red; topic 3) exhibit a seasonal pattern, with peaks corresponding to notable hurricanes while coverage of tribal issues (teal; topic 7) has increased over time.

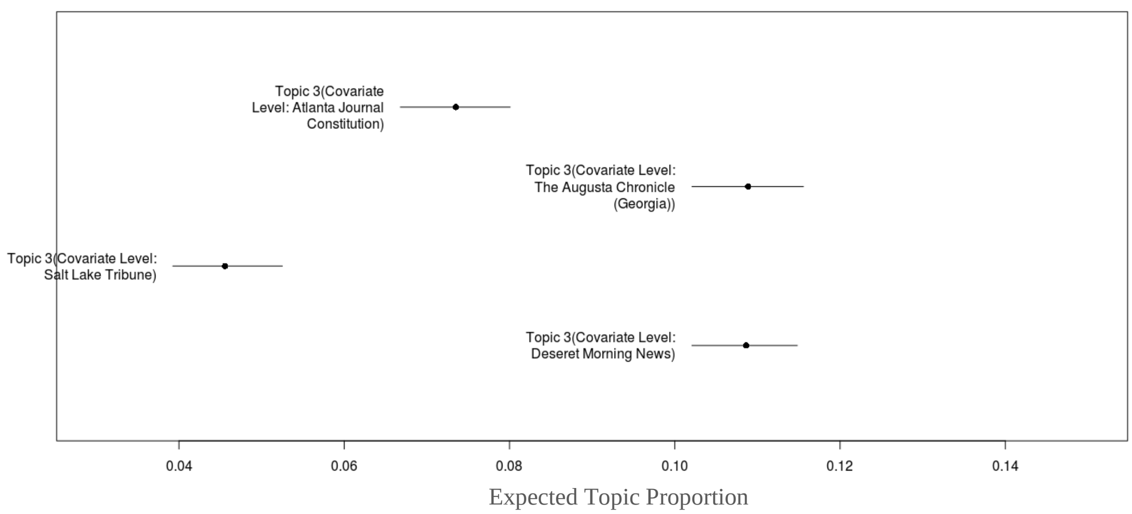

Figure 7. Differences in proportion of topic 3 coverage. The Augusta Chronicle and Deseret Morning News cover hurricanes (topic 3) more than their counterparts in the same states.

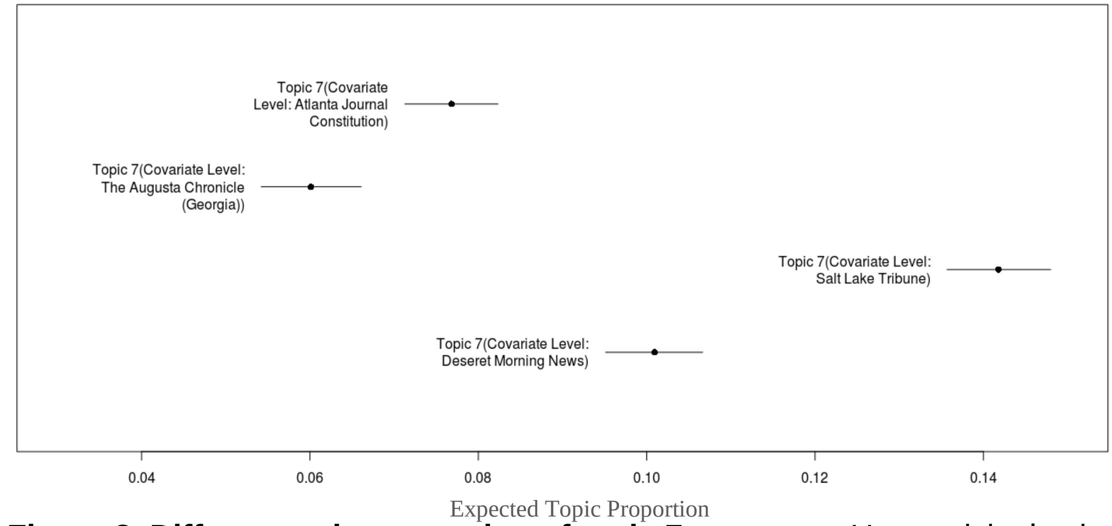

Figure 8. Differences in proportion of topic 7 coverage. Unsurprisingly, the Salt Lake Tribune and Deseret Morning News cover tribal issues (topic 7) more than the two newspapers in Georgia.

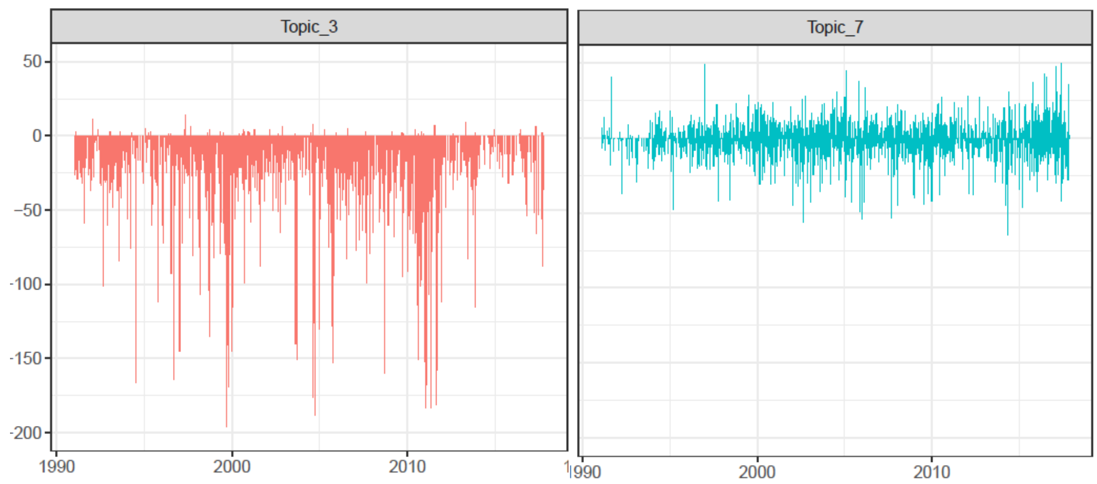

Figure 9. Differences in net sentiment scores for topics 3 and 7. Hurricane (topic 3) coverage is generally negative while tribal issues (topic 7) oscillate between positive and negative sentiments over time.

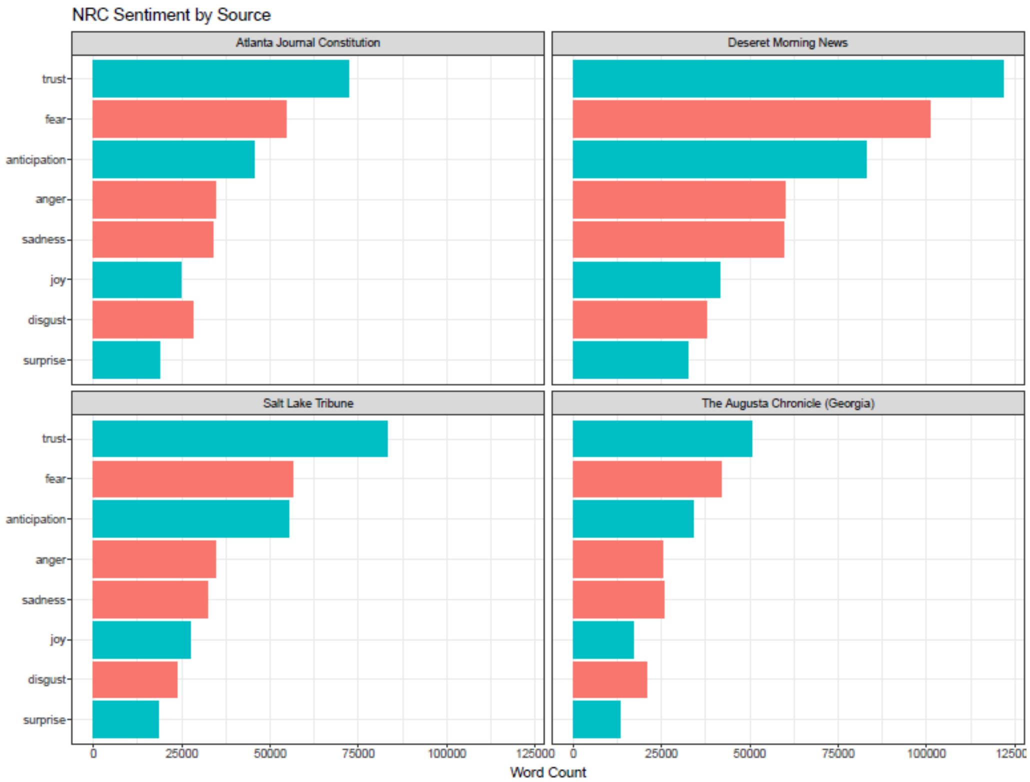

Figure 10. Dominant categories of sentiments across the four sources. Similar categories emerge, with trust and fear, dominating across the newspapers.

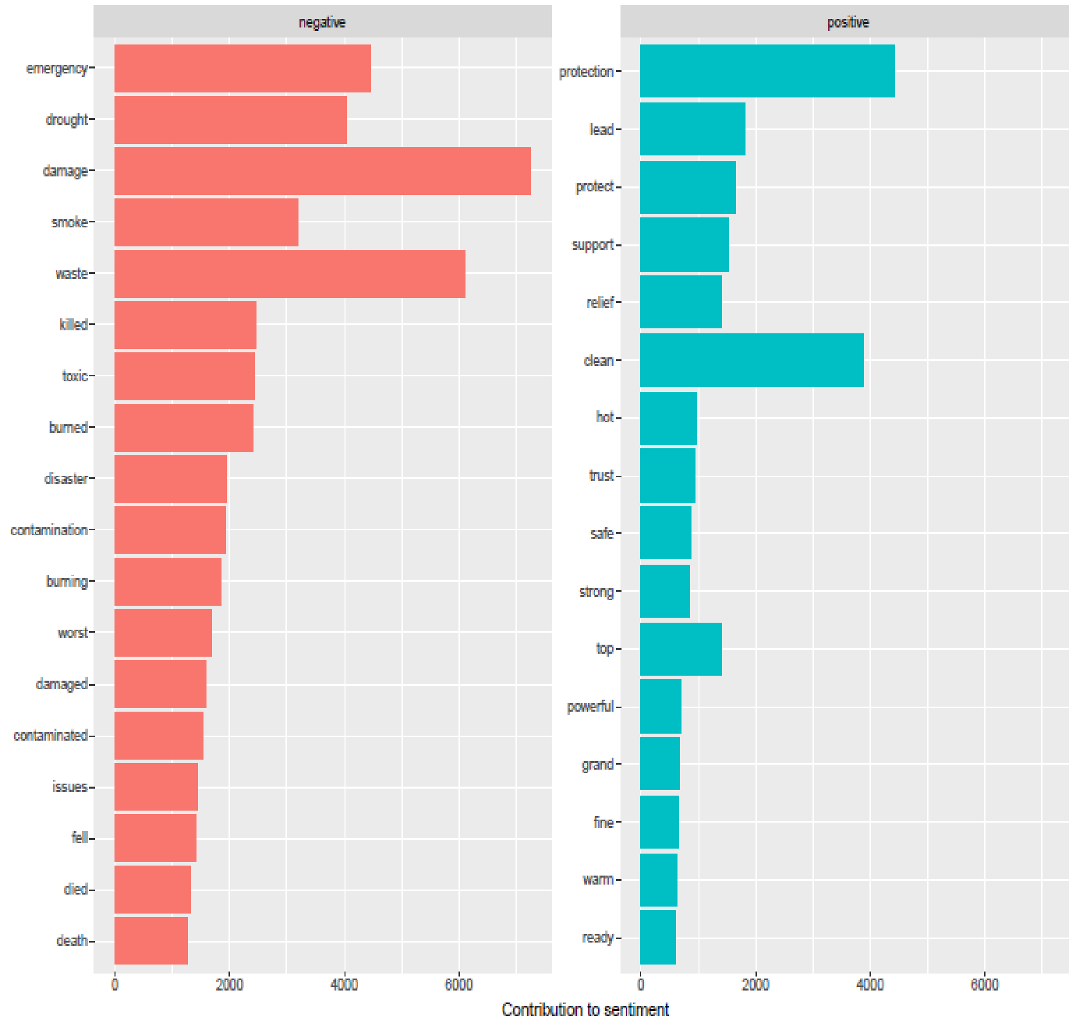

Figure 11. Leading words in the positive and negative sentiment categories. Words like emergencies, droughts, and damage dominate the negative category while words like protection, support, and lead dominate the positive category. The word 'lead', however, could mean "to lead' or "lead in the water" (latter a negative issue), highlighting the limitations with sentiment analysis using bag-of-words.

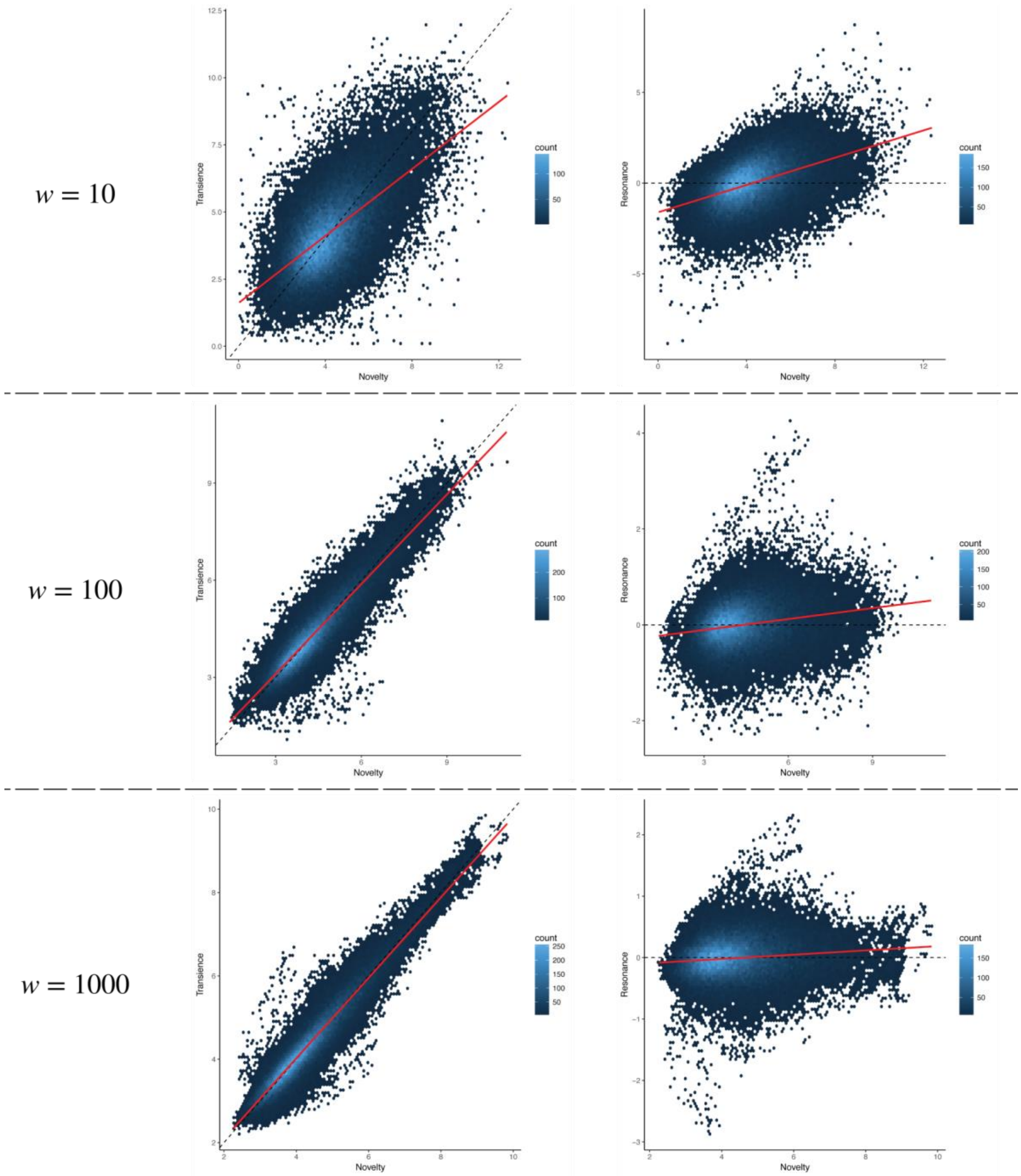

Figure 12. Kullback-Leibler Divergence plots of topic novelty, resonance, and transience. Linear fit lines are in red while the y=x line is in dotted black for novelty vs transience plots and y=0 is in dotted black for novelty vs resonance plots. Window (*w*) corresponds to the size of the window (i.e., number of articles) used in the estimation process both forward and backward in time. Positive slopes of the fitted line (red) in the novelty-resonance relationship indicate that novel topic mixtures persist in future articles.

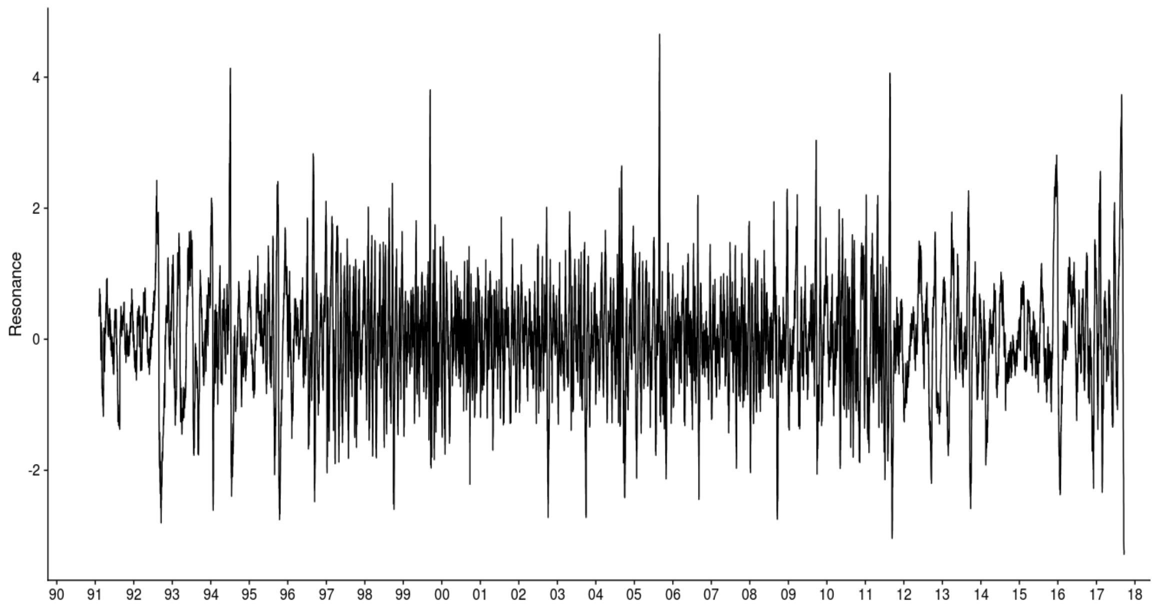

Figure 13. Kullback-Leibler Divergence plot for topic 3. Peak resonance occurred on August 28, 2005, coinciding with Hurricane Katrina's landfall over New Orleans, Louisiana. Window length (*w*) = 100.

4. DISCUSSION

Our results indicate the potential for leveraging NLP methods for gaining insight into narratives about water. In particular, topic modeling approaches can be used to understand the general topics covered in the newspapers as well as associated spatiotemporal patterns. Unsurprisingly, the identified topics predominantly refer to issues such as fires, hurricanes, and waste cleanup (Table 2; Figure 4). Some of the issues are covered in a cyclic manner (e.g., hurricanes) while others exhibit increased coverage over time and in particular regions (e.g., tribal issues in Utah; Figures 6-8). Additional analysis is required to understand the correlations between identified topics (Figure 5) and the impact of the different periods of coverage of newspapers on the generated results. Techniques such as normalizing word counts based on the number of words within the corpus and the length of period coverage might enable a more consistent comparison between the datasets. We could also explore specific words within topics to understand the influence of certain phenomena (e.g., is the word 'drought' primarily restricted to the topic on fires or does it also extend to topics on energy?)

Sentiment analysis provided some insight into the general tone of language used in media coverage. Similar patterns of trust followed by fear emerged across the 4 newspapers analyzed (Figure 10). However, the limitations of this bag-of-words approach quickly became apparent when the sentiments were broken down to specific words (e.g., classification of the word "lead"; Figure 11). A more advanced algorithm such as Linguistic Inquiry and Word Count (LIWC) could provide more refined results, and thus improved understanding of how sentiment varies within and across topics (NASEM, 2018). Alternate approaches such as n-grams (i.e., phrasal analysis) or bi-directional gated recurrent units (GRUs) with learned word embeddings might be required to capture the actual "meaning" of sentences. In addition, a mixed methods approach of manually coding a portion of the articles for tone (following Wei et al., (2017)) to train and analyze the full corpus could also provide additional insights into the actual content covered. The latter approach could enable a rich set of analyses, including but not limited to the prevalence of cooperation versus conflict narratives within media coverage.

The KLD analysis indicates that over large periods of time, new ideas within the articles tend to stick around into the future (Figures 12 and 13). There are a few articles that deviate from this general pattern (i.e., outliers with low novelty and low resonance as well as outliers with medium novelty and high resonance; Figure 12) that require further exploration. Frequency analysis of the topic coverage in conjunction with novelty, resonance, and transience scores could also provide additional insight into observed patterns. A detailed literature review is required to ascertain the novelty of the pointwise KLD approach used for individual topics and to develop an approach for identifying appropriate w values. Generally, the KLD approach provides an interesting avenue for evaluating relationships between policy changes and the persistence of certain narratives within media.

In addition to the aforementioned methods, the analysis needs to be extended to the full corpus. This would help us identify more robust topics that can be used to analyze patterns in all 37 newspapers. With a larger corpus, we can systematically examine variations between newspapers as a function of distances between the cities, as well as other characteristics (e.g., local climate, drinking water source, and other watershed characteristics). Additional metadata that can be explored for patterns include authors of the articles, ideological leanings of the newspaper, and readership size.

The ability of the media to influence the public and subsequently set agendas is well recognized (McCombs & Ghanem, 2001). Wei et al., (2017) show that combining media coverage with local policy changes and local climate conditions can provide insights into the phases of transition and co-evolution of the physical and social systems. In addition to characterizing media narratives, additional data about legislative and other changes can provide insight into the interplay between media coverage and other social actions. Overall, the methods explored in this analysis have the potential to help us understand the evolution of the narratives in relation to local policy changes.

REFERENCES

1. Barron, A. T. J., Huang, J., Spang, R. L., & DeDeo, S. (2018). Individuals, institutions, and innovation in the debates of the French Revolution. *Proceedings of the National Academy of Sciences, 115*(18), 4607-4612. doi:10.1073/pnas.1717729115.
2. Brown, K. P. (2017). Water, water everywhere (or, seeing is believing): The visibility of water supply and the public will for conservation. *Nature and Culture, 12*(3), 219-245. doi:10.3167/nc.2017.120302.
3. de Arruda, H. F., Costa, L. D. F., & Amancio, D. R. (2016). Topic segmentation via community detection in complex networks. *Chaos, 26*. doi:10.1063/1.4954215.
4. Edalat, F. D., & Abdi, M. R. (2018). Concepts and approaches of main water managements. In F. D. Edalat & M. R. Abdi (Eds.), *Adaptive Water Management* (pp. 11-19). New York, NY: Springer.
5. Farrell, J. (2015). Corporate funding and ideological polarization about climate change. *Proceedings of the National Academy of Sciences, 113*(1), 92-97. doi:10.1073/pnas.1509433112.
6. Gao, Y., Xu, Y., & Li, Y. (2013). Pattern-based topic models for information filtering. *2013 IEEE 13th International Conference on Data Mining Workshops* (pp. 921-928). Dallas, TX.
7. Gonzales, P., & Ajami, N. (2017). Social and structural patterns of drought-related water conservation and rebound. *Water Resources Research*, *53*(12), 10619-10634. doi:10.1002/2017WR021852.
8. Gunda, T., Benneyworth, L., & Burchfield, E. (2015). Exploring water indices and associated parameters: A case study approach. *Water Policy, 17*(1), 98-111. doi:10.2166/wp.2014.022.
9. McCombs, M., Ghanem, S.I., 2001. The convergence of agenda setting and framing. In: Reese, S.D., Gandy, O.H., Grant, A.E. (Eds.), *Framing Public Life: Perspectives on Media and Our Understanding of the Social World*. Lawrence Erlbaum Associates, London, pp. 67-81.
10. NASEM (National Academies of Sciences, Engineering, and Medicine). (2018). *Understanding Narratives for National Security: Proceedings of a Workshop.* Washington, DC: The National Academies Press. doi: https://doi.org/10.17226/25119.
11. Quesnel, K. J., & Ajami, N. K. (2017). Changes in water consumption linked to heavy news media coverage of extreme climatic events. *Science Advances*, *3*(10). doi:10.1126/sciadv.1700784.
12. Roberts, M. E., Stewart, B. M., & Airoldi, E. M. (2016). A model of text for experimentation in the social sciences. *Journal of the American Statistical Association, 111*(515), 988-1003. doi:10.1080/01621459.2016.1141684.
13. Roberts, M., Stewart, B., & Tingley, D. (2018). stm: R Package for structural topic models (Version 1.3.3) [Computer software]. Retrieved from https://cran.r-project.org/package=stm.
14. Roberts, M. E., Steward, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science, 58*(4), 1064-1082. doi:10.1111/ajps.12103.

15. Rogers, E. M., & Dearing, J. W. (1988). Agenda-setting research: Where has it been, where is it going? *Communication Yearbook, 11*, 555-594. doi:10.1080/23808985.1988.11678708.
16. Silge, J., & Robinson, D. (2018). Tidytext: Text mining and analysis using tidy data principles in R (Version 0.1.9) [Computer software]. Retrieved from https://cran.r-project.org/package=tidytext.
17. Treuer, G., Koebele, E., Deslatte, A., Ernst, K., Garcia, M., & Manago, K. (2017). A narrative method for analyzing transitions in urban water management: The case of the Miami-Dade Water and Sewer Department. *Water Resources Research*, *53*(1), 891-908. doi:10.1002/2016WR019658.
18. Verón, E. (1997). Esquema para el análisis de la mediatización. *Diálogos de la Comunicación*, 48, 9–17.
19. Wei, J., Wei, Y., & Western, A. (2017). Evolution of the societal value of water resources for economic development versus environmental sustainability in Australia from 1843 to 2011. *Global Environmental Change*, *42*, 82-92. doi:10.1016/j.gloenvha.2016.12.005.
20. Wu, S., Huang, S., Wei, Y., Duffield, C., Tang, W., Zhao, Y., & Zheng, H. (2018). A longitudinal analysis on the perspectives of major world newspapers on the Three Gorges Dam project during 1982–2015. *Water Science and Technology: Water Supply*, *18*(1), 94-107. doi:10.2166/ws.2017.088.

DISTRIBUTION

1	MS0899	Technical Library	9536 (electronic copy)